Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 15 (2011) 2695 – 2699

Procedia Engineering

www.elsevier.com/locate/procedia

Advanced in Control Engineeringand Information Science

Auditory Cryptography Security Algorithm With Audio Shelters

Huan Li [a, b,a*], Zheng Qin[a], Xuanping Zhang[a], Xu Wang[a]

[a]*School of Electronics and Information Engineering, Xi'an Jiaotong University, Shanxi, China*
[b]*Department of Computer Science, Dongguan University of Technology, Guangdong, China*

Abstract

In this paper, auditory cryptography security algorithm with audio shelters is proposed. The meaningful audio watermarking is pretreated to high-fidelity binary audio, and the binary audio is encrypted to *n* cryptographic audios by (*k*, *n*) threshold scheme. Less than *k* of the cryptographic audios give no information, only synchronized playing *k* or more than *k* of the audios the original can be heard directly. The *n* cryptographic audios are embedded in the corresponding *n* shelter audios which are pretreated by high-dimensional matrix transformation. Experiments show that the proposed algorithm has strong practicability, high security and robustness in enduring common attacks.

© 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of [CEIS 2011]

Keywords: auditory cryptography; audio shelter; (k, n) threshold; audio watermarking

1. Introduction

The common audio watermarking algorithms have the following deficiencies: the embedded watermarking signal is meaningless; the watermarking is directly embedded in host audio signal; once the embedded watermark is proposed, the original will be exposed [1-4]. To address these problems, auditory cryptography security algorithm with audio shelters is proposed. The meaningful audio watermarking is pretreated to high-fidelity binary audio, and the binary audio is encrypted to *n* cryptographic audios. By synchronized playing *k* or more than *k* of the audios, the original can be heard directly. The *n* cryptographic audios are embedded in the corresponding *n* shelter audios. Experiments show that the proposed algorithm has strong practicability, high security and robustness in enduring common attacks.

* Corresponding author. Tel.:0086-0769-22861226;
E-mail address: lihuan@dgut.edu.cn.

1877-7058 © 2011 Published by Elsevier Ltd.
doi:10.1016/j.proeng.2011.08.507

2. The related works

2.1. Auditory cryptography (k, n) threshold scheme

Based on binary halftone pretreatment [5], *k* out of *n* Auditory Cryptography (AC) algorithm based on auditory properties of human being is proposed. Firstly, a meaningful audio watermarking is transformed to binary digital audio. Then, using the basic matrix to encrypt fast and safely, and synchronized playing the shares to decrypt directly through Human Auditory System (HAS). Besides, less than *k* of the shares give no information about the original, only synchronized playing *k* or more than *k* of the shares the original can be heard.

Input: the high-fidelity binary audio watermarking w, the length of w is l;

Output: n cryptographic digital audios w_1, w_2, …, and w_n;

Step1: construct basic matrix $B0_{n\times m}$ and $B1_{n\times m}$;

Step2: for any value S_i of w:

if S_i=0, then:

1) using matrix $B0_{n\times m}$;
2) two random position scrambling on row vector and column vector, the scrambled matrix is called extended set of encryption $C0_{n\times m}$;
3) n lines of $C0_{n\times m}$ correspond to n cipher-text sequences;
4) get n sharing values S_{i1}, S_{i2}, …, S_{in};

else if S_i=1, then:

1) using matrix $B1_{n\times m}$;
2) two random position scrambling on row vector and column vector, the scrambled matrix is called extended set of encryption $C1_{n\times m}$;
3) n lines of $C0_{n\times m}$ correspond to n cipher-text sequences;
4) get n sharing values of S_{i1}, S_{i2}, …, S_{in};

Step3: connect all S_{i1} to a digital audio w_1, the same method, get w_2, …, and w_n;

Step4: a plaintext audio w becomes n cryptographic digital audios w_1, w_2, …, and w_n, and the length from l to $m*l$;

2.2. Matrix transformation

Definition 1: 3-dimentional matrix transformation: For a given $m\times m\times m$ (m^3) cube matrix, let $x^3=(x_1, x_2, x_3)$ denote the element coordinate in this cube square matrix and $y^3= (y_1, y_2, y_3)$ be the mapped element coordinate. The following is the 3-dimentional matrix transformation. Where $\boldsymbol{A}_{3\times3}$ is a 3×3 integer coefficients matrix and the determinant of $\boldsymbol{A}_{3\times3}$ ($|\boldsymbol{A}_{3\times3}|$) is a relative prime of m.

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = A_{3\times3} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \bmod m \qquad (1)$$

Definition 2: High dimensional matrix transformation: for a given $m\times m\times\ldots\times m$ (m^n) super cube space, let $x^n=(x_1, x_2,\ldots, x_n)$ denote the element coordinate in this super cube space and $y^n= (y_1, y_2,\ldots,y_n)$ be the mapped element coordinate. The following is the high dimensional matrix transformation. Where $\boldsymbol{A}_{n\times n}$ is an $n\times n$ integer coefficients matrix and the determinant of $\boldsymbol{A}_{n\times n}$ ($|\boldsymbol{A}_{n\times n}|$) is a relative prime of m.

$$\begin{bmatrix} y1 \\ y2 \\ \mathrm{M} \\ yn \end{bmatrix} = A_{n\times n} \begin{bmatrix} x1 \\ x2 \\ \mathrm{M} \\ xn \end{bmatrix} \bmod m \qquad (2)$$

2.3. Evaluation formulas

Some formulas are used to evaluate the scrambling performance. Let N denote the length of audio, T and T' denote the two different audio sequences, and D denote the maximum of the original audio T. These formulas are defined as follows:

Definition 3: BER (Bit Error Rate)

$$BER = \frac{100}{N}\sum_{i=0}^{N-1}\begin{cases}1, T'(i) \neq T(i)\\ 0, T'(i) = T(i)\end{cases} \quad (3)$$

Definition 4: SNR (Signal-to-Noise Ratio)

$$SNR = 10 \bullet \log_{10} \frac{\sum_{i=0}^{N-1} T^2(i)}{\sum_{i=0}^{N-1}\left(T'(i) - T(i)\right)^2} \quad (4)$$

Definition 5: SIM (Angle Cosine)

$$SIM = \frac{\sum_{i=0}^{N-1}\left|T'(i) * T(i)\right|}{\sqrt{\sum_{i=0}^{N-1} T'(i)^2 * \sum_{i=0}^{N-1} T(i)^2}} \quad (5)$$

3. The realization procedure of algorithm

3.1. The frame diagram of algorithm

The frame diagram of algorithm is given in Fig.1. The meaningful audio watermarking is pretreated to a high-fidelity binary audio by the binary halftone pretreatment algorithm [5]. The (k, n) threshold scheme is given in chapter 2.1. The n cryptographic audios $\{w_1, w_2, \ldots, w_n\}$ *are* gotten. The corresponding n shelter audios $\{A_1, A_2, \ldots, A_n\}$ need three steps: 1) dimension increasing; 2) high dimensional matrix transformation[6,7]; 3) dimension decreasing. In this section, the watermarking embedding and watermarking extracting schemes are given.

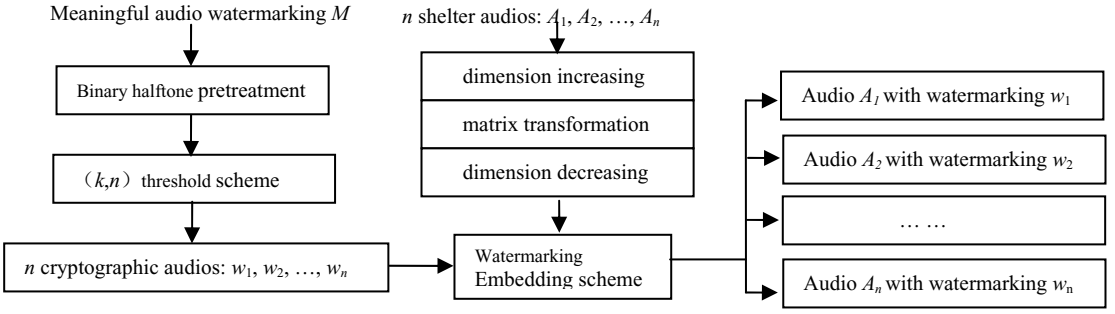


Fig.1. The frame diagram of the proposed algorithm

3.1. Embedding and Extracting Scheme

The n cryptographic audios $\{w_1, w_2, \ldots, w_n\}$ watermarking are separately embedded in the average statistical of corresponding n processed shelter audio signals $A_1, A_2, \ldots, A_n$ sampling values.

Definition 6: Watermark Embedding

Let L denote the length of processed shelter audio signal A_i, which was divided to K segments, and each segment has R audio sample points. Let T denote the pre-processed cryptographic audio

watermarking sequence w_i and $B(i) \in [0, 1]$, $1 \le i \le k$; Let $A_i(j)$ denote the audio point in i segment, $A_i'(j)$ denote the embedded sample point, and S is the max of the values that satisfy the Eq.6:

$$A_i'(j) = \begin{cases} A_i(j) - A_i(j) \bmod S + 3S/4, if B(i) = 1; \\ A_i(j) - A_i(j) \bmod S + S/4, if B(i) = 0; \end{cases} \quad (6)$$

Definition 7: Watermark extracting

According to the idea of maximum likelihood decoding, we extract the watermark. Let $B'(s)$ denote the extracted watermark, $\rho(B', B)$ denote the correlation function:

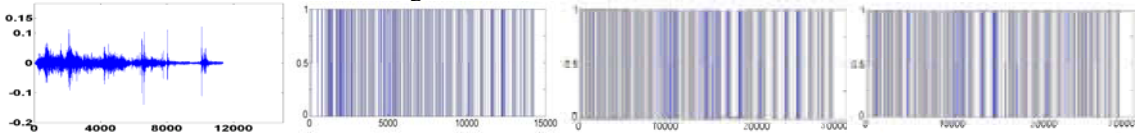
$$\rho(B', B) = \sum_{i=1}^{k} B'(i), B(i) \quad (7)$$

If exits $\rho(B', B) \ge T$, T denote threshold, we think the watermark is extracted.

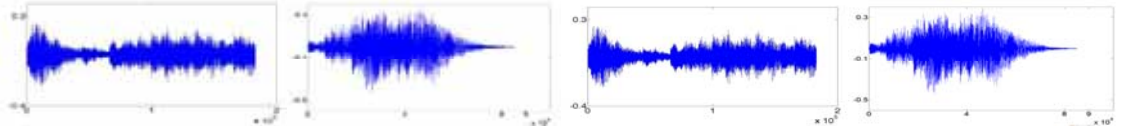
Only when k or more than k of cryptographic audios are extracted, the original audio watermarking can be heard directly

4. Experiment Results and Analysis

The performance of the proposed algorithm is tested in this section. We use the 'start.wav' as audio watermarking sample M (Fig.2a). And by the binary halftone pretreatment algorithm, the corresponding meaningful binary audio watermarking is shown in Fig.2b. According to the chapter 2.1, the (2, 2) threshold scheme is used and two cryptographic audios $\{w_1, w_2\}$ are gotten (Fig 2c, 2d). The two shelter audios 'close.wav' and 'ding.wav' are denoted by A_1 and A_2 (Fig 2e, 2f). The audio A_1 with watermarking w_1 and the Audio A_2 with watermarking w_2 are shown in Fig.2g and Fig.2h.



(a) the audio watermarking 'start.wav' (b) binary audio watermarking (c) cryptographic audio w_1 (d) cryptographic audio w_2



(e) shelter audio 'close.wav' (f)shelter audio 'ding.wav' (g) Audio A_1 with watermarking w_1 (h)Audio A_2 with watermarking w_2

Fig.2. the related audio wave plots

In Table2, we compared the *BER*, *SNR* and *SIM* between the original shelter audios and the audio with watermarking. From Table 2, we can see that the *BER* is lower than 4%, the *SNR* is higher than 30 and the *SIM* is higher than 0.97. So the shelter audios with watermarking are similar with the original shelter audios, and the audios sound hardly any noise. So the proposed scheme has a wonderful embedding performance.

Table 1. watermarking embedding performance of the shelter audios

Parameters		performance		
original shelter audio *A*	audio *A* with watermarking *w'*	*BER* (%)	*SNR* (*db*)	*SIM*
Fig.3e	Fig.3g	3.249	37.54	0.985
Fig.3f	Fig.3h	3.198	38.75	0.977

Typical signal manipulations on shelter audio with watermarking, for instance, resampling, Gaussian noise addition, low pass filtering and mp3 compression, are adopted to attack the watermarked signal. In Table 2, we evaluated the BER, SNR and SIM between the extracted audio watermarking and the original cryptographic watermarking. From Table 2, we can see that the BER value is lower than 9%, the SNR is

higher than 30, and the SIM is higher than 0.83, so the similarity is much high. By playing 2 watermarking audios, the original watermarking audio 'Start.wav' is clearly heard. So the experiment shows that the proposed scheme has a wonderful extracting performance.

Table 2. watermarking extracting performances

Extracted watermarks			Attack Parameters and extracting performances			
cryptographic Watermark	Extracted Watermark	Parameter	Resample 22050hz	gaussian white noise noise SNR=50	lowpass filtering 16khz	MP3 Compression
W_1	W'_1	BER(%)	5.611	6. 453	5.226	7. 425
		SNR(db)	31.25	33.14	37.11	35.05
		SIM	0.965	0.951	0.885	0.863
W_2	W'_2	BER(%)	5.867	5.981	5.515	7.927
		SNR(db)	34.12	34.51	35.82	31.17
		SIM	0.934	0.918	0.857	0.878

5. Conclusion

This paper proposed auditory cryptography security algorithm with audio shelters. The algorithm has following advantages: the embedded watermarking is meaningful audio signal; by (k, n) threshold scheme, the audio watermarking signal is encrypted to n cryptographic audios. Less than k of the cryptographic audios give no information, only synchronized playing k or more than k of the audios the original can be heard directly. The n cryptographic audios are separately embedded in the average statistical of corresponding n processed shelter audio signals. The experiments show that the proposed algorithm has high security and strong robustness.

6. References

[1] Chen Ming, Zhang Ru, Liu Fanfan, Niu Xinxin, Yang Yixian. Audio steganography by quantization index modulation in the DCT domain Audio steganography by quantization index modulation in the DCT domain. Journal on Communications, 30 (8); 2009, p. 105-111.

[2] Xie Chun-hui, Cheng Yi-min, Wang Yun-lu, Chen Yang-kun. Estimation of Secret Message in Audio Based on Statistic Characteristics. Journal of Electronics & Information Technology. 31(6); 2009, p. 1341-1344.

[3]WenQuan,Wang Shuxun,Nian Guijun. Audio Watermarking in DCT Domain: Algorithm and Measurement of Imperceptibility, ACTA Electronica Sinica, 35(9); 2007, p.1702-1705.

[4] Tan Liang, Wu Bo, Liu Zhen, Zhou Mingtian. An Audio Information Hiding Algorithm with High Capacity Which Based on Chaotic and Wavelet Transform. ACTA Electronica Sinica. 38(8): 2010, p. 1812-1824.

[5]Wang Xu, Research on the Security Algorithm of Auditory Cryptography Based on Auditory Properties of Human Being, Master Degree Thesis of Xi'an Jiaotong University, China; 2010.

[6]Shao Liping, Qin Zheng, Li Huan， "Scrambling Matrix Generation Algorithm for High Dimensional Image Scrambling Transformation", IEEE International conference on ICIEA, pp. 3-5; 2008.

[7]Huan. Li, Zheng. Qin, Liping. Shao, "Variable Dimension Space Audio Scrambling Algorithm Against MP3 Compression", IEEE International conference on ICA3PP, pp. 328-335; 2009.